DOCUMENT RESUME

ED 258 455 FL 015 083

AUTHOR Gray, Vicky A.
TITLE Evaluation of the Grade Nine French Immersion Program in Fredericton, N.B.
PUB DATE Jan 85
NOTE 31p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Bilingual Students; Comparative Analysis; Foreign Countries; *French; Grade 9; *Immersion Programs; Language Arts; *Language Proficiency; Listening Comprehension; Program Evaluation; Reading Comprehension; Secondary Education; *Second Language Programs; Vocabulary; Writing Skills
IDENTIFIERS New Brunswick

ABSTRACT

The academic achievement in French of two cohorts of grade 9 students of a public school early immersion program in New Brunswick, which has total French-language instruction through grade 4 and includes all students in the program except transfers was compared to that of a group of native French-speaking students at a nearby French-language school. The students were tested on academic achievement in French language arts (including three reading comprehension tests, a writing skills measure, a general language arts test, and an overall French language proficiency test requiring reading and writing skills) and linguistic competence in French (including test of French morphology and vocabulary and three tests of listening comprehension). Results indicate that the program has met its goal of continued improvement in French skills during grades 7-9, with test score improvement on all of the relevant measures. The immersion students' level of performance on academic measures of French language arts ranged from grade 5 to grade 8 in comparison with the francophone students, and in linguistic competence the comparison was with grade 6 or 7. The students performed at one or two grade levels below immersion students in major Canadian cities. An increase in the complexity and variety of French language heard by the students is recommended to improve their skills. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED258455

BEST COPY AVAILABLE

Evaluation of the Grade Nine French Immersion Program in Fredericton, N.B.

Vicky A. Gray

Psychology Department

University of New Brunswick

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Vicky A. Gray

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

January, 1985

FL 015083

This report presents the results of the grade-9 evaluation of the junior high school phase of the early French immersion program in School District No. 26 in Fredericton, N.B. The evaluation is a continuation of a longitudinal study of the Fredericton program begun when the first students entered the program in grade 1.[1]

The focus of the junior high school evaluation has been on the effectiveness of the program in improving French academic and linguistic skills. Questions about the progress of early immersion students in the junior high school years have been raised by some of the studies designed to compare late immersion programs with early immersion programs. Evaluations of the Montreal late immersion program have found very few differences between the performance of early and late immersion students in grades 8 through 11 on tests of listening and reading comprehension, writing and dictation, and speaking (Adiv, 1980; Genesee, 1979). The fact that students beginning late immersion in grade 7 are able to catch up so rapidly with early immersion students, who have the advantage of seven additional years of immersion instruction, suggests that the early immersion students are not progressing as rapidly as they might. Similarly, Genesee and Leblanc (1978) found little improvement between grade 7 and grade 8 of early immersion.

There has been very little specification of appropriate goals for early immersion programs at the junior high school level. The ideal situation would be to specify the interim performance levels which must be achieved at the junior high school level in order to enable the students to meet the long-term goals of the program by the end of grade 12. Although there is not yet enough data to permit this degree of precision, the current research will contribute to the eventual identification of these interim goals.

A minimal goal would be maintenance of the performance levels achieved at the end of the elementary school program. A more desirable goal would be continued improvement in French skills throughout the junior high school program, as measured by increases in test scores. The most rigorous goal is the achievement of native or native-like performance levels, i.e., immersion scores that are similar to those of Francophones of the same age and ability. Progress toward the goal of native-like proficiency would be reflected in decreases in the difference between the scores of immersion students and those of the Francophone comparison students. The grade-9 data will be evaluated in terms of these goals.

The junior high school immersion program in Fredericton is a continuation of the early immersion program which began in grade 1. For the first three years all instruction is in French, with English instruction delayed until grade 4. From grades 4 through 6 approximately 87 percent of school instruction is in French. In grades 7 and 8 more than half of the students are in a 60 percent French program and the remainder are in a 50 percent French program. In grade 9 they received either 50 or 40 percent French instruction.

Method

Subjects

The data from the first two groups of students to enter the Fredericton early immersion program are presented. The students in Stream 1 completed grade 9 in 1982 and the students in Stream 2 completed grade 9 in 1983. Both groups of students have been tested annually since grade 1.

The immersion group consists of the total population of students enrolled in the program at the grade level being tested, with the exception of students who have transferred into the immersion program from the French school or from other immersion programs. No students are omitted because

of learning difficulties or other problems. The students speak English as their first language and almost all come from monolingual English-speaking homes.

The bilingual French comparison group consists of Francophone students in the appropriate grades at the French school in Fredericton. The composition of this group changes slightly from year to year but the majority of the students have been in the comparison group for several years. The evaluations of the elementary years of the Fredericton program also included a unilingual French comparison group from Edmundston, New Brunswick. In the grade-6 evaluation (Gray, 1981) the scores of the bilingual French group were similar to those of the unilingual French group on all of the French measures.

The characteristics of the grade-9 students are presented in Table 3. The immersion group consists of 64 stream-1 students, tested in 1982, and 96 stream-2 students, tested in 1983, for a total of 160 immersion students. The bilingual French group consists of 28 stream-1 students and 34 stream-2 students for a total of 62 bilingual French students.

The mean age was 179.02 months, with no significant difference between the two groups in age. The home backgrounds were measured by parental occupations rated on Duncan's (1961) decile scale. The mean Duncan rating for the bilingual group (8.32) was lower than the rating for the immersion group (9.08). The parental occupation rating of the bilingual group has dropped slightly from their grade-6 rating of 8.90. The intellectual ability measures used in previous evaluations (Gray, 1981) indicated that both the immersion and the bilingual French groups are of above-average intellectual ability, with no differences between the groups in intellectual ability.

Test Battery

All tests were administered in April and May in 1982 and 1983. Group tests were administered in three sessions of 120 to 150 minutes each and the order of tests within a given session was constant across classes. All testers were bilingual francophones.[2]

The tests used in the evaluation are divided into two categories: those measuring academic achievement in French language arts and those measuring linguistic competence in French. The first measures of academic achievement are tests of reading comprehension. The Test de Compréhension de l'Ecrit (Niveau B, 1979 revision) is a measure of reading comprehension developed by the Bilingual Education Project.[3] It has 22 multiple choice questions based on 10 passages from a variety of sources including newspaper articles, a television schedule, a horoscope, a contest advertisement, and a classified advertisement. A second reading comprehension test with a similar format, the Test de l'Ecrit I (1978-79) developed by the Protestant School Board of Greater Montreal (Adiv and Morcos, 1979), was also given. This test consists of six newspaper articles, one anecdote, and a poem, each followed by multiple choice questions, with a total of 20 questions.[4] The majority of the questions on both of these tests measure the student's understanding of the main ideas in the passages.

The third reading comprehension test, the IEA French Reading Test (Population IV S), was developed as part of an international evaluation of achievement in French as a second language (Carroll, 1975).[5] The test consists of 36 multiple choice questions encompassing a wide range of difficulty. Fifteen items involve filling in a blank or completing a sentence. The remainder are based on short passages. In general, this test tends to place a greater stress on understanding of specific

vocabulary items, phrases, and syntactic structures than the tests described above, and the questions are less likely to be answerable on the basis of the general context of the items.

Writing skills in French were measured by the IEA French Writing Test (Population IV).[5] Part I, Grammar, consists of 25 sentences containing blanks. The student must write in the missing word, usually a verb or a preposition, on the basis of the grammatical context of the sentence. Part II, Verbs and Modifiers, consists of six sentences with verbs, determiners, and adjectives underlined. Each of the six sentences is followed by a parallel sentence which has a change of gender, person, or number and contains blanks corresponding to the underlined words of the first sentence. The student must fill in the 16 blanks to produce grammatical agreement.

French language arts were measured by the Test de Rendement en Français, developed by La Commission des Ecoles Catholiques de Montréal for use in its schools.[6] This test is a comprehensive measure of a variety of French language arts skills, such as word knowledge, parts of speech, spelling, vocabulary, logical sequencing, and comprehension of text. The grade-9 students received both the grade-8 version (Secondaire II, 1980-81) to measure improvement since grade 8 and the version that is designed for grade 9 francophone students (Secondaire III, 1981-82).

The final test concerned with academic achievement in French was the Test de Mots à Trouver (Niveau D, 1977) produced by the Bilingual Education Project.[3] Although this test is considered to be an overall measure of French language proficiency it is included with the academic achievement tests because reading and writing skills are required. However, because successful performance requires a general knowledge of the language, the test also measures overall linguistic proficiency. The test consists of a

passage of prose, entitled "Le Canada vu par un Français", from which every seventh word has been replaced by a blank. The students must fill in the blanks, using their understanding of the surrounding context and their knowledge of written French. The exact scoring method, as described by Lapkin and Swain (1977) was used.

The second section of the test battery measures linguistic competence in French. The first test is the French Morphology Test, which is based on Berko's (1958) test of English morphology. It is designed to measure the student's intuitive understanding of the rules of French word formation. Some items require that the student provide a feminine ending to a masculine nonsense noun; others require the selection of masculine or feminine forms of the indefinite article to test for knowledge of gender cues in the endings of nonsense nouns. There are also items in which the student is required to provide past participle endings for three classes of verbs.

French vocabulary was measured by the French translation of the Peabody Picture Vocabulary Test. The translations are the same as those used in the St. Lambert project (Lambert & Tucker, 1972).[7] This is an individual test requiring the student to select from four pictures the one which best depicts the word given by the tester. Grade 9 students received Form A of the test. (Alternate forms of this test have been administered annually since grade 1.)

Three tests were used to measure listening comprehension. The Test de Compréhension Auditive was designed by the Bilingual Education Project to measure the understanding of spoken French in a variety of real-life situations.[3] The students listen to recorded passages from radio broadcasts, weather forecasts, interviews, etc., and then answer multiple-choice questions based on the contents of each passage. Two

versions of the test were administered. Niveau A (1978 revision, formerly called Test de Compréhension Aurale) has been used since grade 5 and was included to measure improvement. However, the data from the first stream of grade 9 indicated that the bilingual French group had reached the ceiling on this test, with over half of the students making a perfect score and almost all of the remainder making only one error. Because it could not provide an accurate measure of differences between the two groups, this version of the test was not administered to the second stream of grade 9. Niveau B (1979 revision), a more difficult version of the test, was administered to both streams.

The third listening comprehension test was the IEA French Listening Test (Population IV S).[5] The test consists of recorded sentences, short conversations, broadcast announcements, and long conversations. Each selection is followed on the tape by one or more multiple-choice questions, with a total of 37 questions covering a wide range of difficulty.

Results

Data analysis

The main data analysis consisted of analyses of variance computed separately for each measure, with the data for streams 1 and 2 combined to permit more general conclusions. Further comparisons were made with the Tukey test at the .05 level of significance.[8] The analysis of variance results for the grade-7 and grade-8 evaluations are presented in Tables 1 and 2. (See Gray, 1982; 1983 for descriptions of these evaluations.) Table 3 presents the grade-9 results from the current evaluation. In addition, two measures that indicate the magnitude of the differences between the immersion and the French groups are presented in Tables 4 and 5. The rationale for these size-of-effect measures is as follows:

At this point in the immersion research, evaluators of some programs are sometimes failing to find statistically significant differences between immersion students and Francophone comparison students, whereas other evaluators continue to find significant differences between the two groups. In the latter case, it would often be useful to know whether the obtained differences are relatively small or relatively large. In addition, when differences between groups persist over several years, it would be useful to know whether the size of the differences is increasing, decreasing, or remaining the same.

The analysis of variance technique for data analysis allows us to establish whether or not differences between groups are statistically "significant". The designation of a difference as "significant" refers to the reliability of the difference, in the sense that a "reliable" difference is one that has a very low probability of having occurred by chance and thus a high probability of occurring again under similar circumstances. However, the fact that a given difference between two groups is statistically significant does not necessarily mean that the difference is an important difference or even a large difference. For example, if the number of subjects is large enough, a very small and for practical purposes trivial difference between groups may achieve statistical significance. Thus, in addition to establishing that a group difference is statistically significant it is often useful to have information about the size of the difference. Two measures which give this information are presented in this report.

The first size-of-effect measure indicates the percentage of immersion students whose scores on a given test are as high as those of the top 50 percent of the French group, or in other words the percentage of immersion students whose scores fall above the 50th percentile or half-way point for

the French scores. (This measure is determined by calculating the number of immersion students whose score falls above the median score for the French group.) A score of 50 percent on this measure indicates that the scores of 50 percent of the immersion students are as high as those of the top half of the Francophone students, and suggests that the two groups are performing very similarly. A score of zero percent indicates that none of the immersion students have scores as high as those of the top half of the French group.

A more widely used size-of-effect measure is Cohen's d statistic which expresses the difference between two group means in relation to the variability between the groups (Cohen, 1969). (The d value is calculated by dividing the difference between the group means by the weighted average of the standard deviations for each group.) Although the designation of a given difference as a small, medium, or large effect is obviously subjective, Cohen gives the following guidelines: A d value of .8 or larger indicates a large effect, which is described as one that is grossly perceptible, such as the difference in height between 13- and 18-year-old girls, or the difference in IQ between holders of the PhD degree and typical college freshmen. A medium effect is one large enough to be visible to the naked eye, such as the difference in height between 14- and 18-year-old girls, and is indicated by a d value of .5. A small effect, such as the difference in height between 15- and 16-year-old girls, would have a d value of .2 (Cohen, 1969).

Grade 9 results

The grade-9 test results are presented in Table 3. The first tests of academic achievement in French are the reading comprehension tests. On the Test de Compréhension de l'Ecrit (Niveau B) the immersion group's score of 14.89 is significantly lower than that of the bilingual French group, but

shows improvement over their score of 13.93 in grade 8. The size of the difference between the immersion and French groups has decreased slightly from a d value of .41 in grade 8 to .30 in grade 9. In terms of approximate grade equivalent relative to Francophone students, the grade-9 immersion score is somewhat lower than the grade-8 bilingual French score of 15.43. Comparing with immersion programs in other cities, the Fredericton immersion group in grade 9 scores at about the same level as a grade-8 Ottawa group reported by Andrew, Lapkin & Swain, 1980 (14.89 and 14.75 for Fredericton and Ottawa, respectively).

The results of the Montreal Test de l'Ecrit I, administered for the first time in grade 9, are similar. The immersion group scores lower than the bilingual French group, and slightly lower than a group of grade-9 early immersion students in Montreal (13.87 vs 14.64 for Fredericton and Montreal, respectively) (Adiv & Morcos, 1979).

On the IEA French Reading Test the immersion students score lower than the bilingual French group. The grade-9 immersion score is somewhat lower than the score obtained by a group of grade-7 bilingual French students in Fredericton (20.54 vs 22.86). Because this test was not administered before grade 9, there are no measures of improvement. There are also no appropriate comparisons with programs in other cities.

The immersion students appear to perform less well, relative to the French group, on the IEA Reading Test than on the Test de Compréhension de l'Ecrit and the Test de l'Ecrit I. A comparison of the size-of-effect data for the three tests (Table 4) indicates that the difference between the immersion group and the French group is much larger for the IEA French Reading Test than for the other two tests. For example, only 4.48 percent of the immersion students score as high as the top 50 percent of the French group on the IEA test, as compared with 36.18 percent on the Test de

Compréhension de l'Ecrit. This difference in the performance of the immersion students may be a function of the requirements of the tests. The first two tests tend to stress comprehension of the main ideas in the passages, and the meaning of crucial words can often be determined from surrounding context. The IEA test, on the other hand, places greater stress on the precise understanding of specific words, phrases, and grammatical structures, and many of the items offer little contextual support. The fact that the latter test is more difficult is consistent with the observation that immersion students tend to perform better in those areas of communicative competence that do not rely heavily on grammatical knowledge (Harley, 1984).

Writing skills in French were measured by the IEA French Writing Test. As in grades 6 and 8 the immersion group scores at a lower level on both Part I and Part II than the French group. The total score of the immersion group has improved since grade 8 (from 22.05 to 24.36) and now falls above the score of the grade-5 bilingual French students (22.87) but below the score of the grade-6 French students (27.56). The difference between the immersion group and the French group in grade 9 is quite large ($d = 1.24$) and shows no evidence of having decreased since this test was first given in grade 6 ($d = 1.17$) (Table 5). The immersion students' difficulty with this test is probably because of the heavy stress on the grammatical aspects of writing in Parts I and II of the test.

The two levels of the Test de Français provide overall measures of proficiency in French language arts. The grade-8 version (Secondaire II) was administered again to measure improvement. As in grade 8 the immersion students score lower than the bilingual French students, but their grade-9 score of 27.97 is higher than their grade-8 score of 24.69, demonstrating an improvement in their performance. There is, however, no change in the

size of the difference between the immersion group and the bilingual French group; the size-of-effect score in grade 8 (d = .52) is the same as in grade 9 (d = .53). The immersion score at the end of grade 9 (27.97) is roughly equivalent to the score of the bilingual French group at the end of grade 8 (27.92), suggesting an approximate grade equivalent of grade 8 on this measure of French language arts.

On the grade-9 version of this test (Secondaire III) the immersion students score lower than the French group. The mean score of the immersion group falls at the bottom of the fifth stanine range (40th to 59th percentile) and the mean bilingual French score falls at the bottom of the sixth stanine range (60th to 76th percentile) on the grade-9 norms for French-speaking students in Montreal.[9] In comparison with other immersion programs the mean stanine score[7] of the Fredericton immersion group (4.72) is lower than the mean stanine score of a group of Montreal early immersion students in the same grade (5.47) (Genesee & Morcos, 1978).

The final test of French language arts is the Test de Mots à Trouver, an academic measure of overall second language proficiency. As in previous years, the immersion students score lower than the bilingual French students, but have improved from a mean score of 18.27 in grade 8 to their current score of 19.35. However, the magnitude of the difference between the immersion and French groups has not decreased; the grade-9 d value of 1.01 is slightly larger than the grade-8 value of .91. In terms of approximate grade equivalent the grade-9 immersion score falls above the grade-6 bilingual French score of 18.65 but below the grade-7 bilingual French score of 21.02. Comparing with other programs, the grade-9 Fredericton students, with a score of 19.35, are performing at about the same level as grade-8 students in the Carleton/Ottawa area, whose scores

range between 19.3 and 19.9 (Swain & Lapkin, 1982; Morrison, Bonyun, Pawley, & Walsh, 1980).

Swain and Lapkin (1982) have suggested that early immersion students achieve native-like levels of performance on the Test de Mots à Trouver by grade 8. In contrast, in Fredericton the difference between early immersion and French comparison students on this test is still quite large at the end of grade 9 (d = 1.01), with only 14.42 percent of the immersion students scoring as high as the top 50 percent of the French group. Table 6 below shows the scores of the groups cited by Swain and Lapkin (1982, page 42) and the comparable data for the Fredericton groups in grades 8 and 9. These data confirm that immersion students in Fredericton are about a year behind Ottawa immersion students; the immersion group in Fredericton scored lower than the Ottawa-area group in grade 8, but by grade 9 they were performing about as well as the grade-8 Ottawa group. The French comparison students in Fredericton, on the other hand, score much higher than the Ottawa French students; both the grade-8 and the grade-9 scores of the Fredericton group, as well as the grade-7 score of 21.02, are higher than the grade-8 score of the Ottawa group. This suggests that the Francophone score used by Swain and Lapkin is an underestimation of Francophone performance, and that their statement about native-like levels of performance on this test is probably premature.

Table 6. Mean scores on the Test de Mots à Trouver, Niveau D (max = 41)

	Immersion	Francophone
Ottawa area, Grade 8	19.9*	19.6**
Fredericton, Grade 8	18.3	22.1
Fredericton, Grade 9	19.3	23.6

* Ottawa-Carleton Cohort 1 (Swain & Lapkin, 1982)
** Ottawa Roman Catholic School Board (Edwards, Colletta, Fu & McCarrey, 1979)

The second section of the test battery measures linguistic competence in French. On the French Morphology Test the immersion students continue to score lower than the bilingual French group, but the size of the difference is quite small (d = .10) and 35 percent of the immersion students perform as well as the top 50 percent of the French group. In addition, as shown on Table 5, the difference between the two groups has decreased considerably since grade 6. The French group has shown relatively small improvements in their score since the end of grade 4, and appears to be reaching the ceiling for this test. (Our expectation has been that Francophone students would master the rules covered by this test at about this age.) The immersion group, on the other hand, has continued to improve and is beginning to approach Francophone performance levels. The immersion grade-9 score of 14.66 is approximately equivalent to the grade-7 and grade-8 scores (14.42 and 14.59, respectively) of the bilingual French group.

The fact that the immersion students have continued to improve on this test during junior high school is encouraging because the type of internalized, intuitive understanding of the French language measured by this test is probably a major contributor to ultimate levels of French proficiency. Studies of the English version of this test suggest that improvement in this aspect of language knowledge is a function of exposure to the language rather than explicit teaching of rules (Berko, 1958; Gray &

Cameron, 1980).

On the French vocabulary test, the translation of the Peabody Picture Vocabulary Test, the immersion group has improved since grade 7, with mean scores of 99.04 and 108.17 in grades 7 and 9, respectively. (Grade 7 was the most recent year in which Form A was administered.) However, they continue to score lower than the bilingual French group. In terms of approximate grade equivalents, the grade-9 immersion score is somewhat higher than the grade-7 bilingual French score (108.17 vs 103.20). Although a comparison of the size-of-effect scores since grade 6 (Table 5) suggests some decrease in the size of the difference between the two groups, the apparent decrease may be because some Francophone students are scoring at the top of the test range.

The first of the listening comprehension tests is Niveau A of the Test de Compréhension Auditive, which was administered to the first stream of grade 9 students only. The immersion students scored lower than the bilingual French students on this test, and show little improvement over their grade-8 performance (13.73 vs 13.82 for stream 1 in grades 8 and 9, respectively). The immersion students in grade 9 continue to score at the same level as grade-6 French students whose score was 13.48. Since the students have improved on the more difficult version of this test, discussed below, the absence of improvement is not in itself a concern.

On Niveau B of the Test de Compréhension Auditive the immersion students score lower than the bilingual French students, but their grade-9 score of 13.26 is an improvement over their grade-8 score of 10.01. Although there is still a large difference between the immersion and French groups, the d value has decreased from 1.55 in grade 8 to 1.12 in grade 9 suggesting that the difference may be decreasing. The immersion students at the end of grade 9 score lower than grade-7 Francophone students giving

an approximate grade-equivalent somewhat above grade 6. In comparison with students in other cities, the Fredericton students score lower than grade-8 immersion students in the Carleton/Ottawa area (13.28 vs 14.95 for Fredericton and Carleton/Ottawa, respectively) and at about the same level as grade-7 immersion students in Toronto (13.68) (Swain & Lapkin, 1982).

In a summary of immersion evaluation results, Swain and Lapkin (1982, page 54) conclude that the listening comprehension skills of early French immersion students develop to near native-like levels by grade 8, citing as evidence the fact that the scores of Ottawa-area immersion students on Niveau B of the Test de Compréhension Auditive were very similar to those of Ottawa Francophone students. In contrast, at the end of grade 9 Fredericton immersion students continue to score considerably lower than their bilingual French comparison group; the d value on Niveau B is 1.12, indicating a large effect, and only 7.87 percent of the immersion students perform as well as the top 50 percent of the French students.

Table 7 below compares the Fredericton data with the Ottawa-area data cited by Swain and Lapkin. The Ottawa estimate of grade-8 Francophone performance is consistent with the Fredericton data for grade 8, and there is little difference between the scores of the Ottawa immersion students and those of the two Francophone groups. Thus the conclusion of near-native performance levels at grade 8 on this measure of listening competence is reasonable for the Ottawa-area immersion students. However, both the grade-8 and the grade-9 immersion students in Fredericton score lower than those in Ottawa. Fredericton immersion students are about two years behind immersion students in other programs; the Fredericton grade-8 score of 10.01 is slightly lower than the grade-6 score of 11.11 obtained by a group of Ottawa-area students. Similarly, the Fredericton grade-9 score of 13.28 is roughly comparable to the grade-7 score of 13.68 obtained

by a Toronto group (Andrew, et al., 1980). Taken as a whole, these data suggest a weakness in listening comprehension skills for the Fredericton immersion students.

Table 7. Mean scores on the Test de Compréhension Auditive, Niveau B (max =22)

	Immersion	Francophone
Ottawa area, Grade 8	15.0*	14.5**
Fredericton, Grade 8	10.0	14.8
Fredericton, Grade 9	13.3	17.3

* Ottawa-Carleton Cohort 1 (Swain & Lapkin, 1982)
** Ottawa Roman Catholic School Board (Edwards, et al., 1979)

The results of the final test of listening comprehesion, the IEA French Listening Test , confirm that there is a large difference between the immersion and French groups in listening skills. The immersion group scores lower than the bilingual French group, and the d value of 1.54 indicates that this is a very large effect. Only 5.24 percent of the immersion students score as high as the top 50 percent of the French students. The grade-9 immersion score of 22.48 is roughly equivalent to the grade-6 score of 22.16 for grade-6 bilingual French students.

Summary and General Conclusions

The Fredericton program has met the goal of continued improvement in French skills during grades 7, 8, and 9. The immersion students showed improvement on all of the tests that had been administered the previous year in grade 8,[10] with the exception of Niveau A of the Test de Compréhension Auditive which was not appropriate for students at this level.

At the end of grade 9 the performance of the immersion students on the academic measures of French language arts ranges between the grade-5 and the grade-8 level for bilingual French students in Fredericton. On the

measures of French linguistic competence their performance is roughly comparable to that of grade 6 or grade 7 bilingual French students. These approximate grade-equivalents represent improvements of one-half to one and one-half years since the end of grade 8. However, it must be remembered that the grade-equivalent approximations are quite inexact and depend heavily on individual test characteristics.

The performance of the Fredericton students can also be compared with that of immersion students in other programs. On the academic measures for which comparisons are available,[11] the difference is in the vicinity of one year; that is, Fredericton grade-9 students perform about as well as grade-8 immersion students in Montreal, Ottawa/Carleton, or Toronto on the measures of reading comprehension and language arts. However, in the area of listening comprehension the difference is about two years; Fredericton students in grade 9 are roughly comparable to grade 7 students in Toronto. Similarly, Fredericton students in grade 8 performed at about the same level as grade 6 students in the Ottawa area. The two-year difference in listening comprehension again suggests that the Fredericton immersion students are weak in listening comprehension.

The most obvious explanation for the one-year difference in reading and language arts achievement levels is the fact that the Fredericton program begins in grade 1 whereas the other programs begin a year earlier in kindergarten. On the other hand, the Fredericton program has a higher concentration of French after grade 4 than most other programs, which might be expected to compensate for the later start. The fact that the higher concentration of French in the later years of the Fredericton program does not entirely compensate for the absence of kindergarten suggests that both time of exposure and intensity of exposure are important. In particular, it may be that the opportunity to develop a facility in listening

comprehension during kindergarten before the onset of formal instruction in reading and writing in grade 1 may promote more rapid progress in all areas of French skills.

Although immersion performance improved in terms of increases in test scores, the evidence for improvement in performance relative to the French comparison group is mixed; the immersion students continue to score lower than the French comparison students on all measures, and on most tests the difference between the scores of the two groups is large. Table 5 compares the size-of-effect scores for the tests that have been administered for two or more years.

On the Test de Compréhension de l'Ecrit, which is not a difficult test for students at this level, and Niveau B of the Test de Compréhension Auditive there is some evidence of a decrease in the difference between the scores of the two groups. On the French Morphology Test the difference between the immersion group and the French group has decreased steadily over the last few years. This is because the French students are, as expected, beginning to master the rules covered by this test and their scores are showing little improvement, whereas the immersion scores are continuing to improve.

There is no evidence of improvement relative to the French group on the other tests and in a few cases, such as the Test de Mots à Trouver, there is some suggestion that the difference between the two groups is increasing rather than decreasing. Although improvement relative to French comparison students would be a very positive finding, it may also be an unrealistic expectation. More data on achievement levels and magnitude of group differences for the post-elementary years of immersion programs in Fredericton and in other cities is necessary before realistic expectations for achievement relative to Francophone populations can be specified.

Nevertheless, some investigators have reported native-like performance levels, i.e., immersion scores that are similar to those of Francophones of the same age, in the areas of reading and listening comprehension. Lapkin and Swain (1984, p. 51) in a summary of immersion evaluation results conclude that "For early-immersion students, the receptive skills are clearly native-like by the end of elementary school." Although this conclusion may be premature, at least for immersion programs as a whole, it is also clear from the results discussed above that the Ottawa/Carleton and Toronto students evaluated by Lapkin and Swain are closer to native-like performance in the receptive skills of reading and listening than are Fredericton students. The difference is particularly marked for listening comprehension. There is insufficient information to say whether the Fredericton results or the Ottawa-area results are more representative of immersion programs across the country, but the data do indicate that there is room for considerable improvement in listening comprehension in the Fredericton program, and it is recommended that this improvement should be considered a priority.

Listening comprehension is widely believed to be fundamental to other aspects of language, and improvements in listening comprehension would probably be accompanied by improvements in other skill areas. In the development of first-language skills the ability to comprehend various aspects of language almost always preceeds the ability to produce these aspects and many theorists (e.g., Krashen, 1984) assume that the same is true for second language learning. The difficulty encountered in immersion programs is that the primary exposure to spoken French is in the classroom, and in many cases the classroom language is generated almost entirely by the teacher. Further, the language of the teacher is often limited to the language of classroom management. Immersion students quickly acquire

enough facility with French to understand the language of the classroom, but unless special steps are taken the ordinary language of the classroom is probably lacking in the complexity and variety that are necessary to stimulate continued improvement. Thus, the most promising way to produce improvement is to increase the complexity and variety of the language heard by the students by encouraging the use of radio, television, and films both in the classroom and at home, by guest speakers and field trips, by extra-curricular activities involving contact with native speakers, and so on.

Footnotes

1. The evaluation was funded by Secretary of State and School District No. 26 in Fredericton, New Brunswick.

2. The Coordinator of Testing was Juanita Mureika in 1982 and Deborah Johnston in 1983.

3. Permission to use these tests was granted by the Bilingual Education Project, Ontario Institute for Studies in Education.

4. Permission to use this test was granted by Dr. Ellen Adiv of the Protestant School Board of Greater Montreal.

5. Permission to use these tests was granted by Dr. John B. Carroll and the International Association for the Evaluation of Educational Achievement.

6. Permission to use these tests was granted by La Commission des Ecoles Catholiques de Montreal.

7. The French translations of the Peabody Picture Vocabulary Test were provided by Dr. W.E. Lambert.

8. The statistical analyses presented in this report were computed using SPSS (Statistical Package for the Social Sciences), Version 9. This program package is maintained by the University of New Brunswick Computing Centre. Marni Turnbull and Murray Linton conducted the statistical analyses for this report.

9. The stanine equivalent of the mean score for a particular group should not be confused with the mean stanine score for a group. The former is simply the stanine value that corresponds to the group mean score, whereas the latter is calculated by converting each student's score to its stanine equivalent and then calculating the mean of these stanine equivalents. This conversion to stanine scores allows for a rough comparison across different test versions.

10. All of the tests except the Test de l'Ecrit I, the IEA French Reading Test, and the IEA French Listening Test were administered in grades 8 and 9 (grades 7 and 9 for Form A of the vocabulary test).

11. Comparisons with other programs are available for the Test de Compréhension de l'Ecrit (OISE), the Test de l'Ecrit I (PSBGM), the grade-9 version of the Test de Français (Secondaire III), and the Test de Mots à Trouver.

References

Adiv, E. A comparison of second language performance in two alternative French immersion programs. Montreal: Protestant School Board of Greater Montreal, 1980.

Adiv, E., and Morcos, C. A comparison of three alternative French immersion programs at the grade 9 level. Montreal: Protestant School Board of Greater Montreal, 1979.

Andrew, C.M., Lapkin, S., and Swain, M. Report on the 1979 evaluation of a French immersion program and an extended French program in the Toronto Board of Education, grades 5-8. Toronto: Ontario Institute for Studies in Education, 1980.

Berko, J. The child's learning of English morphology. Word, 1958, 14, 150-177.

Carroll, J.B. The teaching of French as a foreign language in eight countries. New York: Wiley, 1975.

Cohen, J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1969.

Duncan, O.D. A socioeconomic index for all occupations. In A.J. Reiss, Jr. (Ed.) Occupations and social status. Glencoe, Ill.: Free Press, 1961.

Edwards, H.P., Colletta, S., Fu, L., and McCarrey, H.A. Evaluation of the Federally and Provincially Funded Extensions of the Second Language Programs in the Schools of the Ottawa Roman Catholic Separate School Board: Annual Report 1978-1979. Toronto: Ministry of Education, Ontario, 1979.

Genesee, F. A comparison of early and late immersion programs. Montreal: McGill University, 1979.

Genesee, F. and Leblanc, M. A comparative evaluation of three alternative French immersion programs: Grades 7 and 8. Instructional Services Department, Protestant School Board of Greater Montreal, 1978.

Genesee, F. and Marcos, C. A comparative evaluation of three alternative French immersion programs: Grades 8 and 9. Montreal: McGill University, 1978.

Gray, V.A. Evaluation of the grade six French immersion program in Fredericton, N.B. Fredericton, New Brunswick, 1981.

Gray, V.A. Evaluation of the French immersion program in Fredericton, N.B.: Grades seven and eight. Fredericton, New Brunswick, 1982.

Gray, V.A. Evaluation of the French Immersion Program in Fredericton, N.B.: Grades eight and nine. Fredericton, New Brunswick, 1983.

Gray, V.A. and Camerson, C.A. Longitudinal development of English morphology in French immersion children. Applied Psycholinguistics, 1980, 1, 171-181.

Harley, B. How good is their French? Language and Society, 1984, No. 12, 55-60.

Krashen, S. D. Immersion: why it works and what it has taught us. Language and Society, 1984, No. 12, 61-64.

Lambert, W.E. and Tucker, G.R. Bilingual education of children. Rowley, Mass.: Newbury House Publishers, 1972.

Lapkin, S. and Swain, M. The use of English and French cloze tests in a bilingual education program evaluation: Validity and error analysis. Language Learning, 1977, 27, 279-314.

Lapkin, S. and Swain, M. Research update. Language and Society, 1984, No. 12, 48-54.

Morrison, F., Bonyun, R., Pawley, C., and Walsh, M. French proficiency and general progress: Students in elementary core French programs, 1973-1980, and in immersion and bilingual programs, grades 8, 10 and 12, 1980. Toronto: Ministry of Education, Ontario, 1980.

Swain, M. and Lapkin, S. Evaluating Bilingual Education: A Canadian Case Study. Clevedon, Avon: Multilingual Matters, 1982.

Table 1

Test Results of Grade 7 Immersion and Bilingual French Groups

	Immersion (N=162)	Bilingual French (N=69)	F ratio	df	p
Age in months	155.03	155.71	1.17	1,229	ns[b]
Parental occupation					
Duncan decile score	9.04	8.49	4.55	1,227	.03
Test de Français (max = 35)	19.84	22.04	6.75	1,218	.01
(Secondaire I)	(4.13)[a]	(4.85)[a]	9.42	1,218	.001
Peabody Picture Vocabulary Test (French translation, Form A)	99.04	103.20	2.75	1,229	(.10)
French Morphology Test (max = 18)	12.30	14.42	39.60	1,229	.0001

[a] Mean of stanine equivalents calculated from test norms for Francophone students.

[b] Not significantly different.

Table 2

Test Results of Grade 8 Immersion and Bilingual French Groups

	Immersion (N=162)	Bilingual French (N=61)	F ratio	df	p
Age in months	166.72	167.21	.63	1,221	ns[b]
Parental occupation Duncan decile scale	9.08	8.36	7.70	1,219	.01
School Achievement					
Test de Compréhension de l'Ecrit (Niveau B, max = 22)	13.93	15.43	7.28	1,209	.01
IEA French Writing Test					
Part I (max = 25)	12.77	17.72	68.18	1,213	.0001
Part II (max = 16)	9.28	12.12	36.14	1,213	.0001
Test de Français (max = 40)	24.69	27.92	11.41	1,210	.001
(Secondaire II)	(4.91)[a]	(5.83)[a]	11.86	1,210	.001
Test de Français (max = 35)	22.41	25.88	10.81	1,189	.001
(Secondaire I)	(4.84)[a]	(6.18)[a]	18.19	1,189	.0001
Test de mots à Trouver (Niveau D, max = 41)	18.27	22.10	34.52	1,209	.0001
Language Competence					
French Morphology Test (max = 18)	13.72	14.59	6.99	1,219	.01
Peabody Picture Vocabulary Test (French translation, Form B)	98.59	108.84	15.81	1,219	.0001
Test de Compréhension Auditive (Niveau A, max = 17)	13.52	15.65	41.69	1,213	.0001
Test de Compréhension Auditive (Niveau B, max = 22)	10.01	14.80	102.35	1,207	.0001

[a] Mean of stanine equivalents calculated from test norms for Francophone students.

[b] Not significantly different.

Table 3

Results of Grade 5 Immersion and Bilingual French Groups

	Immersion (N=160)	Bilingual French (N=62)	F ratio	df	p
Age in months	178.82	179.52	1.25	1,220	ns[a]
Parental occupation Duncan decile scale	9.08	8.32	7.62	1,217	.01
School Achievement					
Test de Compréhension de l'Ecrit (Niveau B, max = 22)	14.89	15.98	3.97	1,209	.05
Test de l'Ecrit I PSBGM, max = 20)	13.87	15.42	12.18	1,209	.001
IEA French Reading Test (max = 36)	20.54	27.64	70.48	1,212	.0001
IEA French Writing Test					
Part I (max = 25)	14.58	18.63	49.52	1,208	.0001
Part II (max = 16)	9.78	13.27	61.78	1,208	.0001
Test de Français (max = 35)	19.51	22.83	16.75	1,215	.0001
(Secondaire III)	(4.72)[b]	(5.82)[b]	19.42	1,215	.0001
Test de Français (max = 40)	27.97	31.05	11.60	1,207	.001
(Secondaire II)	(5.83)[b]	(6.78)[b]	12.08	1,207	.001
Test de Mots à Trouver (Niveau D, max = 41)	19.35	23.59	43.75	1,212	.0001
Language Competence					
French Morphology Test (max = 18)	14.66	15.29	4.07	1,220	.05
Peabody Picture Vocabulary Test (French translation, Form A)	108.17	116.86	13.90	1,220	.0001
Test de Compréhension Auditive[c] (Niveau A, max = 17)	13.82[c]	16.11[c]	22.22	1,87	.0001
Test de Compréhension Auditive (Niveau B, max = 22)	13.28	17.34	66.96	1,211	.0001
IEA French Listening Test (max = 37)	22.48	31.00	100.68	1,213	.0001

[a] Not significantly different

[b] Mean of stanine equivalents calculated from test norms for Francophone students

[c] Stream 1 only, Immersion n = 62, Bilingual French n = 27

Table 4

Size-of-effect Values for Grade 6 Tests

	Percentage above median[a]	d-value[d]
School Achievement		
Test de Compréhension de l'Ecrit (Niveau B, max = 22)	36.18	.30
Test de l'Ecrit I (PSBGM, max = 20)	29.44	.53
IEA French Reading Test (max = 36)	4.48	1.27
IEA French Writing Test		
Part I (max = 25)	16.76	1.08
Part II (max = 16)	7.78	1.20
Test de Français (max = 35) (Sec. III)	18.04	.62
Stanine[b]	20.38	.67
Test de Français (max = 40) (Sec. II)	29.58	.52
Stanine[b]	29.22	.53
Test de Mots à Trouver (Niveau D, max = 41)	14.42	1.01
Language Competence		
French Morphology Test (max = 18)	35.00	.10
Peabody Picture Vocabulary Test (French translation, Form A)	27.32	.55
Test de Compréhension Auditive (Niveau A, max = 17)	7.77[c]	1.30[c]
Test de Compréhension Auditive (Niveau B, max = 22)	7.87	1.12
IEA French Listening Test (max = 37)	5.24	1.53

[a] Percentage of immersion students who score higher than the median score of the bilingual French students

[b] Calculated from stanine equivalents

[c] Stream 1 only

[d] Cohen's d; see text for explanation

Table 5

Size-of-effect Values for Selected Tests

	Percentage above median[a]	d Value[b]
Test de Compréhension de l'Ecrit (Niv. B)		
Grade 8	29.08	.41
Grade 9	36.18	.30
IEA French Writing Test (total score)		
Grade 6	17.70	1.17
Grade 8	14.84	1.22
Grade 9	6.29	1.24
Test de Français (stanine)		
Grade 6	22.49	.73
Grade 7		.45
Grade 8	35.46	.52
Grade 9	20.38	.67
Test de Mots à Trouver (Niv. B)		
Grade 6	21.91	.83
Grade 8	16.99	.91
Grade 9	14.42	1.01
French Morphology Test		
Grade 6	20.38	.94
Grade 7	17.92	.91
Grade 8	34.98	.40
Grade 9	35.00	.10
Peabody Picture Vocabulary Test (French)		
Grade 6	21.56	.74
Grade 7	43.43	.24
Grade 8	21.25	.59
Grade 9	27.92	.55
Test de Compréhension Auditive (Niv. A)		
Grade 6	10.05	1.25
Grade 8	12.95	.99
Grade 9	7.77	1.10
Test de Compréhension Auditive (Niv. B)		
Grade 8	5.42	1.55
Grade 9	7.87	1.12

[a] Percentage of immersion students who score higher than the median score of the bilingual students

[b] Cohen's d; see text for explanation.